

Thermodynamics of Emergence: Langton's Ant Meets Boltzmann

Heiko Hamann, Thomas Schmickl, Karl Crailsheim

Artificial Life Lab of the Department of Zoology, University of Graz, Universitätsplatz 2, A-8010 Graz, Austria
heiko.hamann@uni-graz.at

Abstract—The current definitions of emergence have no effects in the context of artificial life that could convincingly be called ‘constructive’. They are rather descriptive labels or tests. In order to get towards recipes of generating emergence we need to know systemic characteristics that help during the design phase of artificial life systems and worlds. In this paper, we develop and discuss five hypotheses that are not meant to be irrevocable but rather thought-provoking. We introduce two modeling approaches for Langton's ant to clarify these hypotheses. Then we discuss general properties of systems, such as (ir-)reversibility, dependence on initial states, computation, discreteness, and undecidable properties of system states.

I. INTRODUCTION

Emergence as a concept in the context of artificial life can be seen as both a possibly eminent concept or a vague concept of questionable relevance to the everyday live of an artificial life researcher. Many positions seem to be viable because a fully clarifying theory is pending. Does emergence actually exist? Is there an appropriate definition for it? In case it exists, where and under which conditions is it observed? Is it relevant for the field of artificial life or even beneficial for it?

In this paper, we argue that more research needs to be done before we could dare to give a definition and begin to understand it. The authors have the impression that emergence is an extensive concept. Any approach towards a complete theory of emergence will be inherently interdisciplinary and a too narrow conception would distort the signification of emergence. Hence, studies to formalize it are too early at this stage. First, some rather fundamental questions need to be answered. For example, a set of examples needs to be found that can safely be considered emergent without logical errors. The maximal reward of this patience could be a new theory with an immense impact.

The intention of this paper is a humble attempt to be a thought-provoking impulse. The main substance we are able to promise is that we juxtapose several concepts from other fields that might be worth investigating in depth. We try to point up that these concepts are relevant to any useful concept of emergence. In this article, we argue that the concept of emergence is intrinsically complex and interdisciplinary. Hence, any complete theory of emergence will be complex and interdisciplinary as well. Here, we take the view point of physics and computation. In a case study, we apply methods of thermodynamics and statistical mechanics to Langton's ant—a system that is commonly considered to be emergent [1]. Based on this example system we discuss the concepts related

to (ir-)reversibility, undecidable properties of system states, and infinity. Considering these concepts we find potential thermodynamic and computational approaches to emergence. We develop and discuss five hypotheses that might prove to be relevant to emergence:

- 1) Micro-macro problems in modeling indicate emergence.
- 2) The seeming ‘emergence’ of reversible systems exclusively relies on untypical (i.e., improbable) initial states.
- 3) Emergence is based on irreducible computations and therefore only dissipative systems can be emergent.
- 4) Discrete systems are not emergent.
- 5) Emergence (in reversible systems) depends on undecidable differences between initial states.

In this paper, we cannot give an extensive discussion of definitions of emergence but only a concise summary of definitions that we take as the starting point for this work. We also leave out any discussion of profound philosophical aspects and restrict ourselves to physical and computational issues only.

II. EMERGENCE

Defining emergence in the context of artificial life and multi-agent systems is a challenging research topic. Possibly, Holland is right in his pessimism stating [2]: ‘It is unlikely that a topic as complicated as emergence will submit meekly to a concise definition, and I have no such definition to offer.’ The original idea of emergence was to define a theory that describes the ‘concept of genuinely new kinds of properties produced by nature that cannot be reduced’ [3]. Usually, two levels are introduced in definitions of emergence: a macroscopic level and a microscopic level. The probably most relevant definition is ‘weak emergence’ or the definition by the ‘only means is simulation’-argument by [4], [5]:

A macro-state P of S with micro-dynamic D is weakly emergent iff P can be derived from D and S 's external conditions but only by simulation. [...] for P to be weakly emergent, what matters is that there is a derivation of P from D and S 's external conditions and any such derivation is a simulation.

The authors consider the definition of emergence based on the simulation argument as the one that is most robust to objections. A similar definition is reported in [6]: ‘A true emergent phenomenon is one for which the optimal means of prediction is simulation.’ Also the definition by Standish [7]

is related: ‘An emergent phenomenon is simply one that is described by atomic concepts available in the macrolanguage, but cannot be so described in the microlanguage.’

An alternative but fuzzier concept is the ‘not explicitly defined’-argument. See for example [8]:

In other words, we can speak of emergent individual behavior if the resulting robot behavior was not explicitly programmed in any of its functional blocks and arises from interactions among them [...] and with the environment.

However, without a formal theory the distinction between explicitly and implicitly defined behaviors is rather subjective. A second alternative is the concept of ‘intrinsic emergence’ by Crutchfield who argues: ‘Emergence is meaningless unless it is defined within the context of processes themselves’ [9]. The idea is to overcome the requirement of an external observer [10], [9]. Such a definition is intriguing but leads to the subtle issue of novelty as stated in the paper title: ‘Is anything ever new?’ [9], which is currently a rather informal concept. The question whether a system’s macroscopic state is really new or whether it is defined on the micro-level already before the instantiation stays unanswered. As discussed by Stepney et al. [11]: How could anything new actually emerge from something that is already existent? However, Bickhard and Campbell [12] argue that this question is ill-posed.

Other definitions that rely, for example, on the Aristotelian ‘the sum is more than its parts’-argument seem to have a shortcoming. For example, we refer to Kubík [13]:

Basic emergence then refers to a property of the system that can be produced by interactions of its agents (components) with each other and with the environment and cannot be produced by summing behaviors of individual agents in the environment.

Such a definition seems to include, for example, almost every man-made machine. However, it excludes systems such as termites building clusters of wood chips [14] or ants building clusters of corpses [15]. In these scenarios a single agent would produce the same outcome and would have the same functionality. Therefore, these scenarios should not be defined to be emergent according to Kubík. Still, we will focus mainly on the ‘only means is simulation’-argument in this paper.

III. A CASE STUDY: LANGTON’S ANT

At first, we present a case study, to which we will refer mainly throughout the general discussion of irreversibility, infinity, and undecidable properties of initial states. As our example we choose Langton’s ant (LA) [1]. There is a lot of literature that extends Langton’s original ant and that deeply investigates characteristics of it. See for example [16], [17], [18]. The actual and only relevance here is that LA might be considered to be emergent because, for example, for a given initial configuration we cannot know for sure whether it will show the typical highway pattern. Hence, the typical approach would be to simulate the ant and to search for a highway pattern. For the ant’s world we choose a finite, toroidal grid.

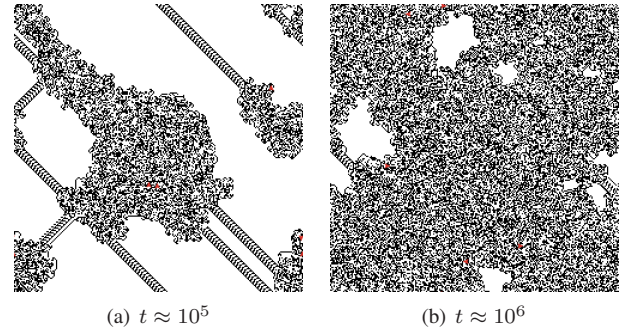


Fig. 1. Patterns generated by five co-existing Langton’s ants.

Hence, the investigated system is neither infinite nor is it irreversible as we argue in the following. Whether this is good or bad news for the potential emergence of LA will be discussed below.

At first, we discuss which property of this system can be modeled with methods of statistical mechanics. The ant’s highway pattern corresponds to a periodic pattern in the sequence of right and left turns which, in turn, corresponds to a periodic pattern in the sequence of colors (white or black) of discovered grid sites. This sequence of colors corresponds to a certain ratio of white to black sites in both the set of all visited grid sites and accordingly scaled over time and space in the fully available grid. Hence, we use the ratio of white sites to all grid sites as an analogon of the thermodynamic entropy. The number of white sites is denoted by $W(t)$, the total number grid sites by N , and the number of black sites in the grid is hence given by $N - W(t)$. The ratio of white sites is given by $S(t) = W(t)/N$. The state of highest entropy would be indicated by $S = 0.5$ (ignoring artificial states such as a half all black, half all white grid), states of lowest entropy are $S = 0$ and $S = 1$. The importance of highways is diminished in this work because the grid is finite and any highway will be disrupted when the ant approaches sites that had been visited before. In addition, we will investigate multiple co-existing ants on one grid. We argue below that this is a valid approach in the context of analyzing the ant’s potential for emergence. Fig. 1 shows typical grid configurations at two time steps for five ants (randomly positioned) in a 240×240 toroidal grid that was initially 100% white.

In the following we derive a macroscopic model for $S(t)$ similar to the statistical mechanics by Boltzmann and we demonstrate that the approach of ensemble averages seems to be infeasible here.

A. Stoßzahlansatz

We have a grid of $\sqrt{N} \times \sqrt{N}$ grid sites and there are M LAs on the grid. We want to find a model for the ratio of white sites

$$S(t) = W(t)/N. \quad (1)$$

We extend this by a distinction between sites that have not yet been visited by any ant (unvisited sites u) and sites that have already been passed by at least one ant (passed sites p).

In a second distinction we further subdivide these numbers into unvisited black sites u_b , passed black sites p_b and for white sites respectively. Rewriting eq. 1 accordingly yields $S(t)N = u_w(t) + p_w(t)$. Introducing $\Delta(t) = p_w(t) - p_b(t)$ we can write this without using p_w and p_b explicitly

$$S(t)N = u_w(t) + \frac{1}{2}(N - u_w(t) - u_b(t) + \Delta(t)). \quad (2)$$

Now we focus on finding a good representation of Δ . With b and w we denote the number of black and white sites respectively that will be visited by an odd number of ants in the next step (i.e., they will switch their state). Hence, we get

$$\Delta(t+1) = \Delta(t) + w(t) - b(t). \quad (3)$$

Both b and w are composed of both unvisited and passed sites. We get

$$w(t) = \alpha(t)p_w(t) + \beta(t)u_w(t), \quad (4)$$

for the ratio α of passed white sites that will be visited by an odd number of ants, the ratio β of unvisited white sites that will be visited by an odd number of ants and

$$b(t) = \gamma(t)p_b(t) + \delta(t)u_b, \quad (5)$$

for ratios γ and δ accordingly. We extend eq. 3 and get

$$\begin{aligned} \Delta(t+1) = \Delta(t) - \alpha(t)p_w(t) - \beta(t)u_w(t) \\ + \gamma(t)p_b(t) + \delta(t)u_b(t). \end{aligned} \quad (6)$$

Now we have to ignore the microscopic details. There is no way of treating w and b without considering the trajectories of the ants. Hence, we approximate the real situation with an analogue of Boltzmann's Stoßzahlansatz (molecular chaos) [19], for an easy-to-read introduction see [20]. First, we assume that the ants have visited already many sites. Hence, we assume that the ratio α of passed white sites, that will be visited by an odd number of ants, and the ratio γ of the black counterpart are equal. That is $\forall t : \alpha(t) = \gamma(t)$. In addition, we restrict the model to the special case of $u_b(0) = 0$ here, for an easier notation. Simplifying eq. 6 yields

$$\begin{aligned} \Delta(t+1) = \Delta(t) - \alpha(t)(p_w(t) - p_b(t)) - \beta(t)u_w(t) \\ = (1 - \alpha(t))\Delta(t) - \beta(t)u_w(t) \end{aligned} \quad (7)$$

Second, we assume that the ants need many steps to reach all sites and that the sequence of sites, which are visited by the ants, is uncorrelated concerning their states. Hence, the number of unvisited sites u_w is proportional to the probability of being unvisited for the site that was initially farthest from any ant and the trajectory of LAs has similar properties as a random walk on a lattice. We can model the probability of being unvisited for any site following [21, eq. 9b]

$$u_w(t)/N = c_1 \ln(\sqrt{N_e}) / \ln(N_e) \exp(-t/(A_R N_e \ln(N_e))), \quad (8)$$

for a constant $c_1 = 2$ that is used to scale $u_w(0) = 1$, an effective number of lattice sites N_e , and a constant A_R that depends on the dimension of the grid. The constant for bounded 2-d lattices is $A_R = 0.44$ according to [22]. By

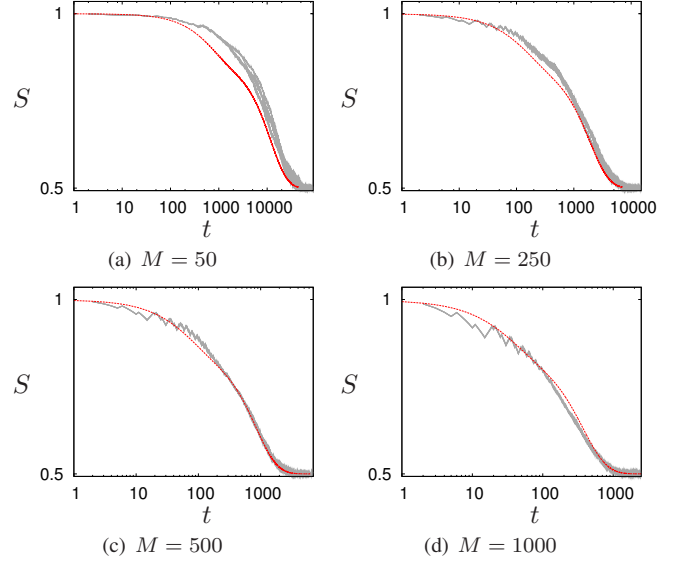


Fig. 2. Plots of the ratio of white sites S as measured in the simulation and by the model with double logarithmic scale (grid size $N = 240^2$, M gives number of ants).

investigating the point process of placing M ants (points) on a $\sqrt{N} \times \sqrt{N}$ lattice we obtain the expected maximal distance between an ant and a lattice site. This distance is used to calculate the number of sites each ant has to explore by subdividing the toroidal grid into bounded squares. The fitted function for the probability of being unvisited of that site that was initially farthest from any ant is given by

$$N_e(M) = \left(1.41374(\sqrt{M/N})^{-1}\right)^2. \quad (9)$$

Finally, we assume that in the average the probability of a passed site being visited by an ant is dependent on the ant density and the density of passed sites in the grid. We neglect the possibility that more than one ant visits a site at the same time since the probability that two ants are at the same site is less than 0.001 for all considered settings. The probability of a site to get a first visit by the ant depends on the density of unvisited sites and the ant density. Hence, we get

$$\alpha(t) = (N - u_w(t))M/N^2, \quad (10)$$

$$\beta(t) = c_2 u_w(t)M/N^2, \quad (11)$$

with a constant $c_2 = 1/(2\pi)$ that scales the area of passed sites to the outline of the area of unvisited sites. By combining eqs. 2, 7, 8, 10, and 11 the macro-model is completed now. In Fig. 2 we compare simulation results with the model. For $M < 500$ the true value of S is underestimated but for $M \geq 500$ the model gives a very good prediction.

B. Ensemble Average

Despite these good predictions the Stoßzahlansatz model has shortcomings as discussed below. For example, we see that $S(t) \rightarrow 0.5$ as $t \rightarrow \infty$ due to eq. 8 which does not correspond to the microscopic behavior as shown below (e.g., see Fig. 4).

Therefore, we try to derive an alternative approach. In the following, we demonstrate the difficulty of deriving an ensemble average for the difference between the number of white and the number of black sites $\Delta'(t) = u_w(t) + p_w(t) - p_b(t)$. With $\chi(\mathbf{r}, t)$ we denote the color of the site at \mathbf{r} and time step t , whereas 1 represents a black site and -1 a white site. With $m(\mathbf{r}, t)$ we denote the number of ants that will visit the site at \mathbf{r} . Hence, the dynamics of χ is given by

$$\chi(\mathbf{r}, t+1) = (-1)^{m(\mathbf{r}, t)} \chi(\mathbf{r}, t), \quad (12)$$

(note, $-1^0 = 1$, for m_e even: $-1^{m_e} = 1$, and for m_o odd: $-1^{m_o} = -1$) and we get

$$\begin{aligned} \Delta'(t) &= \sum_{\mathbf{r}} \chi(\mathbf{r}, t) \\ &= \sum_{\mathbf{r}} (-1)^{m(\mathbf{r}, t-1)} \chi(\mathbf{r}, t-1) \\ &= \sum_{\mathbf{r}} (-1)^{m(\mathbf{r}, t-1) + m(\mathbf{r}, t-2) + \dots + m(\mathbf{r}, 0)} \chi(\mathbf{r}, 0). \end{aligned} \quad (13)$$

Taking now the ensemble average yields

$$\langle \Delta'(t) \rangle = \sum_{\mathbf{r}} (-1)^{\langle m(\mathbf{r}, t-1) + m(\mathbf{r}, t-2) + \dots + m(\mathbf{r}, 0) \rangle} \chi(\mathbf{r}, 0). \quad (14)$$

Hence, the problem reduces to evaluating $\langle m(\mathbf{r}, t-1) + m(\mathbf{r}, t-2) + \dots + m(\mathbf{r}, 0) \rangle$. Now we have to address the trajectories of the ants. Say the position of ant k in the grid at time t is given by $\mathbf{A}_k(t)$. Then we get

$$m(\mathbf{r}, t) = \sum_{k=1}^M \delta(\mathbf{A}_k(t), \mathbf{r}), \quad (15)$$

for a 2-d variant of Kronecker's delta (gives 1, if the two parameters are equal; otherwise 0). At this point, we have a problem because we do not want to evaluate the complete ant trajectories $\mathbf{A}_k(t)$. This would be a microscopic model. In case of continuous systems, the usual way out is to find a differential equation for the trajectories $\mathbf{A}_k(t)$ that could be integrated over a small time interval (e.g., see [23], [24]). In case of a discrete system one would have to find a difference equation. However, it seems that there is no way of finding an integrable, closed-form representation of the ant's behavior that would, for example, correctly describe periodic micro-behaviors (cf. Fig. 4). Therefore, this approach seems to fail.

IV. FUNDAMENTAL PROBLEMS IN MICRO-MACRO MODELING

The case study with LA is an example for the fundamental problems in deriving macroscopic models based on microscopic behavior. Systems, that consist of highly correlated micro-entities, are difficult to model macroscopically because microscopic details (e.g., correlations) cannot be ignored or, if ignored, the model gives wrong predictions. This complex of problems is described by the concept of the micro-macro problem as discussed in [24]:

The *micro-macro link* or also called *micro-macro problem* is a technical term in sociology [25],

[26]. This term is based on the concept of micro-interactions (interactions between individual humans, for example, actions that are aimed to change the actions of others) and macro-structures (structural rules defined, for example, by the human society). The micro-macro link describes the mutual influence of the macro-structure to micro-interactions and vice versa. The action of an individual might be influenced or caused by the individual's perception of the macro-structure. This behavior influences, in turn, the macro-structure leading to a closed loop. Due to this loop structure it is difficult to distinguish whether the cause of a considered action is found on the micro- or the macro-level as indicated by the term 'micro-macro problem'.

This rather fuzzy defined problem of separating the micro-level from the macro-level becomes concrete in the derivation of macro-models of multi-agent systems (as discussed below) and in compiling global-to-local programs [27], [28].

One cause of the micro-macro problem might be, as argued in [29], that novelty is generated in a creative step or creative leap in emergent systems. This corresponds to the general problem of deriving macro-models based on micro-models. For example, based on the micro-model 'Langevin equation' (a stochastic differential equation describing, for example, Brownian motion) the macro-model 'Fokker-Planck equation' (a deterministic partial differential equation describing the evolution of a probability density) can be derived; however, in a critical step of the derivation a high number of collisions per time is assumed to be able to usefully average over them [23], [24]. If the mean free path is long, another macro-model, the Boltzmann equation, is used which describes the probability density over space, time, and velocity of a particle in a medium (kinetic theory). However, in its general form it includes the so-called collision term which describes the effects of particle-particle collisions. Ludwig Boltzmann (1844–1906) determined this term applying his famous Stoßzahlansatz (as discussed in the previous section). He only considers collisions of pairs of particles that are assumed to be uncorrelated prior to the collision. This is, however, not generalizable because particles often are correlated prior to the collision. The macro-model used for higher densities (short mean free path) is the Navier-Stokes equation (fluid mechanics). Interestingly the transition from the Boltzmann equation to the Navier-Stokes equation with increasing density is deducible only for special cases. Even the approach of Lattice Boltzmann methods, that is an altogether different approach, shows similar problems for high Mach numbers, that is, particles with high velocities and few collisions per time [30], [31]. The persistence of these problems independent of the modeling method and depending on the particle densities (i.e., number of correlations) indicates a generality of the micro-macro problem.

These problems persist, if models, such as the Langevin or the Fokker-Planck equation, are applied to emergent systems of autonomous agents. Always, a creative step is necessary in the derivation of the macro-model because micro-behavior relies

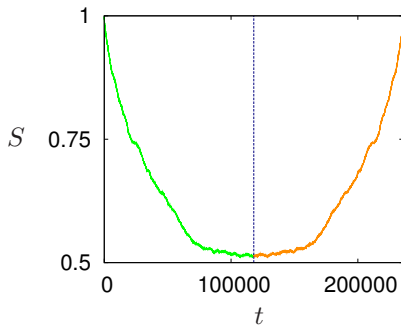


Fig. 3. Reversibility: At $t = 117,937$ the directions (i.e., velocity vectors) of all ants are inverted which corresponds to the inversion of time. The subsequent behavior is an inverted replay of the former behavior.

on macro-structures [24]. It is an open question whether this creative step is necessary due to our ignorance or because this novelty was actually generated by the emergent system (cf. [10], [9]). We call this issue the micro-macro problem which is the challenge of including the correlations of the particles into the macro-model. We claim that the micro-macro problem is a symptom of emergence and should, hence, be included in a theory of emergence.

V. REVERSIBILITY

In this section, we want to resume the discussion of the relation between micro- and macro-models of multi-agent systems. There is a qualitative difference between micro-models and macro-models. For example, the micro-model of gas kinetics (Newtonian mechanics) is time-reversible while the thermodynamic macro-model (Boltzmann equation) is irreversible. Still, Boltzmann claimed to have derived the second law of thermodynamics (macro-model) based on the micro-model via statistical mechanics (Stoßzahlansatz). Today it is widely acknowledged that Boltzmann gave only a proof for the ideal gas limit. How is it possible to derive irreversible macro-rules based on reversible micro-rules? This is discussed in an easy-to-read paper by [20] based on the example of the Kac ring model [32]. The ‘solution’ is to assume that the particles are uncorrelated. This way, the history of the system is fully ignored. The Boltzmann entropy increases monotonically over time (cf. section III-A). Thus, it is irreversible. The derivation of an irreversible model based on a reversible model was only possible by ignoring the correlations between particles (assumption of statistical independence). This asymmetry was criticized by Loschmidt’s reversibility paradox [33].

Now we want to discuss whether reversibility is relevant for emergent systems. Note that Standish [7] considers irreversibility itself an emergent property. It turns out that systems, that are considered to be emergent, show differences concerning reversibility. LA is obviously (although unproven) time-reversible [34]. This is easily shown by running it, for example, until a highway has formed. By inverting the ant’s current direction it will completely remove the highway and return into its initial state. If we keep it running ‘back in time’ it will start another highway which we could remove again by

inverting the direction again. Note that this holds also for a group of coexisting LAs, if their order of execution is kept (see Fig. 3). Thus, in case of LA we have a full time-symmetry and there is no favored direction of time.

A more traditional example (similar to Maxwell’s demon) is a box with gas that is divided into two halves. Say all the gas is in one half of the box – a state of low entropy. Say we remove the barrier in the middle. After some time, the gas will be equally distributed in the box – a state of high entropy. However, in this configuration the particles are highly correlated. We will refer to this state of equally distributed but correlated particles by *Loschmidt configuration* which has a special property: If we invert the velocities of all gas particles they will approach the earlier state of being all in one half of the box again (low entropy) because the system is reversible. The important question is: What can we learn from that in terms of emergence? In case of a reversible system the initial state is of importance and we have to ponder which states are typical. Hence, we should investigate whether a certain initial state or a class of initial states occurs with a reasonable probability in the ‘natural’ context of the system (citing Boltzmann [35]: ‘Aside from a vanishingly small number of special initial states, the most probable states will also occur the most frequently (at least for a very large number of molecules).’). If you choose to start with a Loschmidt configuration it should not be surprising to see an entropy decrease and some kind of pattern formation. For example, one could start with the configuration at time $t = 117,937$ of the experiment shown in Fig. 3. Then the ‘emergent’ phenomenon would be the removal of all black sites (before new black sites are created in a new cycle). Therefore, the emergence of highway formations (or similarly the generation of low entropy states based on Loschmidt configurations) by LA could be considered an artifact of choosing a fully homogeneous initial state which is arbitrarily improbable compared to the number of possible states 2^N .

VI. IRREVERSIBILITY

Now we want to turn to irreversible systems that show emergent behavior. An example from our lab is the BEECLUST algorithm that allows a swarm to aggregate at a maximum of a gradient field although individual agents do not perform a greedy gradient ascent [36]. Controlled by this algorithm two agents will stop when they approach each other, measure the local value of the gradient, and wait for some time proportional to this measurement. Possibly others will join and aggregate in a cluster. For a sufficiently steep gradient the two measurements of two initial agents in such a cluster will be different, thus also their waiting times, and one robot will leave before the other. This behavior causes the irreversibility of the system. Inverting time would mean that the robot, that left later before, would stop without approaching another robot now. This uncaused stopping is an eminent problem of causality. Why should a robot stop without a significant change in its sensor input?

Theorizing about this peculiarity can actually help in understanding emergent systems. In a frictionless system of natural gas or billiard balls, there is no such problem of uncaused stopping with inverted time because collisions have immediate and mutual effects. The autonomous robots controlled by BEECLUST, however, have an ‘autonomy of ignoring’ such effects. The staying robot stays unaffected when the other one leaves. This behavior causes an asymmetry in the robot-robot-interactions and in causality. In contrast to the billiard balls the robots have an inner state that can change without a change in the sensor input. This is the actual cause triggering the leave of the second robot. There is a counter implemented by a variable that is reduced in each cycle until it reaches zero which triggers the end of the waiting phase. The problem with inverted time is that this counter would need to change from zero to one spontaneously.

The above mentioned ‘autonomy of ignoring’ is in fact a process of forgetting information which has special physical properties. In a standard electronic circuit the process of resetting a variable is at least connected to a thermal process, that is, thermodynamics. Therefore, it is important to clarify that the BEECLUST scenario relies on irreversible processes such as diffusion of heat. Hence, we claim that Landauer’s principle [37] might be important in this context, in contrast to the following statement: ‘However, for many so-called self-organizing systems the physical reality is quite detached, i.e., conservation or other thermodynamic laws are not (and need not be) part of the model dynamics, so Landauer’s principles [37] [...] are irrelevant in the general scenario’ [38]. We want to object to this statement as the thermodynamic characteristics that would be ignored by this detachment might be a prominent feature to the theory of emergence.

The distinction between reversible and irreversible computation [39] might be relevant for emergence. Logically reversible computations are defined by a transition function that maps old computational states to new ones by a one-to-one function. Note that in contrast to natural systems, reversible computation models are very unreliable. Consider for example the billiard-ball model [40]. Arbitrarily small perturbations would quickly generate chaotic deviations. In contrast, nature evolved robust computational, emergent systems as, for example, the trail system of ants that is based on the irreversible process of evaporating pheromones. Hence, the question arises whether thermodynamic features such as the irreducibility of diffusion processes are the sine qua non of emergence.

VII. INFINITY

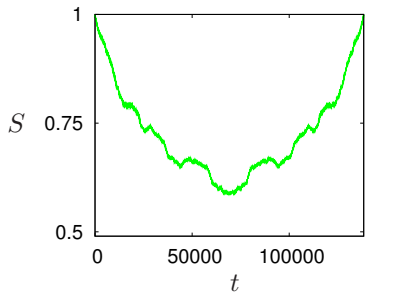
From a computational point of view, there is an essential difference between multi-agent systems defined in discrete or continuous worlds. For example, the set of emergent patterns for a discrete world is recursively enumerable, that is, we could in principle define an algorithm that generates and outputs all patterns ordered by their size without leaving anything out. In case of continuous worlds, this is not possible because the set of real numbers is uncountable, that is, there is no ordering that would not miss infinitely many patterns between

two neighboring patterns (for an entertaining introduction see [41]).

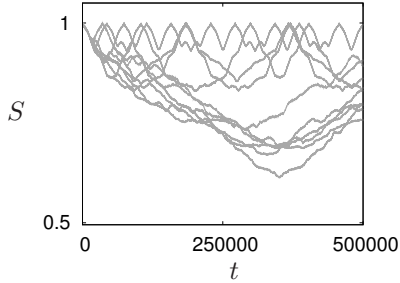
In the following, we argue that emergence can only exist in continuous worlds, however, let us assume for now that it could exist in discrete worlds. We focus on the ‘not explicitly programmed’-argument for a moment. Therefore, one property of emergent systems would be that there is no rule in the system that describes the potentially emergent phenomenon directly. However, for every discrete system and for all of its finite or recurrent patterns (e.g., the ant’s highway) we can derive automatically a trivial macro-system that has a 1-time-step rule that immediately produces the behavior of this pattern. This is done by summarizing the computations of several time steps and several relevant discrete states (e.g., states of grid sites) in macro-rules. For example, in case of the glider pattern (recurrent, size 3×3 , and a period of four steps) in the Game of Life (GOL) we have to summarize 3×3 -packages of the grid as one macro-site and rules summarize four steps in one macro-step. Accordingly the number of possible macro-states increases to 2^9 instead of just two micro-states (dead or alive). In case of LA, this could be done in an extreme by summarizing the initial about 10,000 steps, that are needed before the ant starts the highway on a all-white grid, in one giant macro-step. Thus, neither for the macro-glider nor for the macro-ant-highway one could argue that the behavior is not explicitly programmed. Counter-arguments could be that these macro-systems would explicitly incorporate the complexity that was formerly intrinsic to the system. In addition, the rule system would be very complicated and it would be a pre-computation of the simulation. Still, this algorithm exists and could be used to find ‘emergent’ patterns automatically. Standish argues: ‘An emergent phenomenon is simply one that is described by atomic concepts available in the macrolanguage, but cannot be so described in the microlanguage’ [7]. Is the micro-to-macro-language distinction still relevant, if we can compute the macro-language?

The ‘only means is simulation’-argument is also a weak corrective here because it is difficult to argue why a strictly finite (and in reference to a particular pattern even constant) number of pre-calculations should already count as simulating; or differently put, which calculations qualify as being simulations and which are regular calculations as they would be necessary to calculate, for example, (non-chaotic) planetary motion. We try to find a distinction between mathematical calculations and simulation through a definition of mathematics by Mach [42]: ‘Mathematics is the method of replacing new numerical operations by beforehand carried out, that is, not to be reiterated operations as far as possible and in the most economical way.’¹ We claim that our above mentioned method, hence, qualifies as a mathematical calculation instead of a simulation. Even the definition of weak emergence via explanatory incompressibility [43] does not clearly disqualify this reasoning because at the least the dynamics of the system would be compressed by macro-rules. Finally, the Turing

¹Translation by the authors.



(a) The cycle has length 138,466 and after the first half (69,233) the behavior is reverted and replayed as if time would have been inverted.



(b) 10 samples, cf. Fig. 2, here no log. scale

Fig. 4. Finiteness: As the grid size is limited, the behavior of the ants has to be cyclic ($M = 2$ ants).

Universality of GOL is not an argument because our approach is limited to finite patterns.

Reconsidering our model from section III-A, issues connected to (in-)finite systems are Zermelo's paradox [44] and Poincaré recurrence theorem [45]. Zermelo criticized that Boltzmann's entropy actually cannot decrease forever even in continuous worlds. In LA (on a discrete, finite torus) this can be imitated as shown in Fig. 4. On a finite grid all sequences of grid configurations have to be cyclic. The obtained cycle of length 138,466 (Fig. 4(a)) is far from the theoretical maximum of 2^{240} (ignoring the combinatorics of ant positions). The typical rejection to Zermelo's paradox is that the recurrence time (in continuous worlds) is practically close to infinity (citing Boltzmann: 'it has many trillions of digits' [46]). This is surely not true for discrete (especially, for finite and discrete) systems. Hence, all investigated patterns in the dynamics of discrete systems are cyclic and, even worse, the cycle times are often practically relevant (as shown in Fig. 4). Therefore, the concept of emergence in discrete worlds is questionable.

VIII. UNDECIDABLE PROPERTIES

There is a peculiarity about reversible systems concerning undecidable properties of system states. Imagine we have generated a Loschmidt configuration in the above mentioned box of gas and an additional configuration that is fully random (basically, all velocity vectors point to equally distributed directions and are of random lengths). How could one tell the difference? The only way seems to be by means of simulation. Hence, it could be argued that the correlated particles constitute an emergent property following the 'only

means is simulation'-argument. Note that in agreement with thermodynamics a randomly chosen initial state will with high probability result in dynamics that increases entropy during a reasonable sized time interval while Loschmidt configurations are infrequent. Again the question arises: Does the emergent system generate the unlikely Loschmidt configurations itself or are they used to initialize the system (i.e., they originate from outside of the system)?

Note that deciding whether an initial state results in emergent behavior (i.e., testing for Loschmidt configurations) might be undecidable in general. This concept can be extended in the following way. To prove that emergence is an undecidable property of initial states, a good approach would be to try to prove semi-decidability. One approach could be to show the semi-decidability of emergent systems, that is, a test that answers 'yes' for systems, that are emergent, and that gives no answer for non-emergent systems. The anti-approach to emergence would be to show that only the non-emergence is semi-decidable, that is a test that answers 'no' for systems that are not emergent and that gives no answer for emergent systems. For example, for finite patterns a successful automated construction of a 1-time-step rule would prove the non-emergence of this pattern but that would be needed for all patterns.

We elaborate on this in a little gedankenexperiment. We take a standard example of an undecidable problem: the Post Correspondence Problem (PCP) [47]. An easy conversion (even though a little absurd) to the multi-agent domain could be to make the word pairs embodied objects floating in a medium and showing Brownian motion. We would need an infinite number of them and an infinitely large pool. This example's absurdity is relativized by its similarity to artificial and natural processes, such as the non-biological physical organisms reported in [48] or the DNA replication. Say we chose initial states that represent instances of PCP and we call finished pattern formations (i.e., solutions to PCP) emergent due to their perfection (at least in accordance with the 'not explicitly programmed'-argument). Then it would provably be undecidable whether a given initial state will result in emergent behavior. This shows in addition to the above mentioned Loschmidt configurations that the concept of emergence is related to undecidable properties of system states (see also [49]).

IX. CONCLUSION

We have reported two modeling techniques for LA. The Stoßzahlansatz fails because the grid sites are correlated. Its predicted monotonic entropy increase is incorrect because the system is cyclic. The ensemble approach would require a correct, closed-form approximation of the ant's behavior within small time scales which is unknown. Hence, there is a micro-macro modeling problem that might indicate emergence. We have argued that in reversible systems the selection of an initial state is relevant for emergence. In another hypothesis we stated that irreducible computations might be a sine qua non of emergence. For discrete systems we argued that the recursive

enumeration of ‘emergent’ patterns might allow automatic computations of macro-languages. Finally, we proposed the hypothesis that emergence and undecidable properties of system states might be related and that emergence might depend on undecidable differences of initial states. Future work is to consider each hypothesis, to elaborate on sophisticated definitions of characteristics of emergence, and to develop constructive recipes to design emergent systems.

ACKNOWLEDGMENTS

This work is supported by: EU-IST-FET project ‘SYMBRION’, no. 216342; EU-ICT project ‘REPLICATOR’, no. 216240.

REFERENCES

- [1] C. G. Langton, “Studying artificial life with cellular automata,” *Physica D: Nonlinear Phenomena*, vol. 22, no. 1–3, pp. 120–149, 1986.
- [2] J. H. Holland, *Emergence - From Chaos to Order*. New York: Oxford University Press, 1998.
- [3] R. Schumacher, “Book review: Achim Stephan: Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation. Dresden/München: Dresden University Press, 1999,” *European Journal of Philosophy*, vol. 10, no. 3, pp. 415–419, Dec. 2002.
- [4] M. A. Bedau, “Weak emergence,” *Philosophical Perspectives: Mind, Causation, and World*, vol. 11, pp. 375–399, 1997.
- [5] —, “Downward causation and the autonomy of weak emergence,” *Principia*, vol. 6, no. 1, pp. 5–50, 2002.
- [6] V. Darley, “Emergent phenomena and complexity,” in *Artificial Life IV*, R. Brooks and P. Maes, Eds., 1994, pp. 411–416.
- [7] R. K. Standish, “On complexity and emergence,” *Complexity International*, vol. 9, 2001.
- [8] A. Martinoli, “Swarm intelligence in autonomous collective robotics: From tools to the analysis and synthesis of distributed control strategies,” Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, 1999.
- [9] J. Crutchfield, “Is anything ever new?” in *Complexity: Metaphors, Models, and Reality*, ser. SFI Series in the Sciences of Complexity proceedings, G. Cowan, D. Pines, and D. Melzner, Eds., vol. 19. Reading, MA: Addison-Wesley, 1994, pp. 479–497.
- [10] —, “The calculi of emergence: Computation, dynamics, and induction,” *Pyhsica D*, vol. 75, no. 1–3, pp. 11–54, Aug. 1994.
- [11] S. Stepney, F. Polack, and H. Turner, “Engineering emergence,” in *CEC 2006: 11th IEEE International Conference on Engineering of Complex Computer Systems, Stanford, CA, USA*. Los Alamitos, CA: IEEE Press, Aug. 2006.
- [12] M. H. Bickhard and D. T. Campbell, “Emergence,” in *Downward Causation: Minds, Bodies and Matter*, P. B. Andersen, C. Emmeche, N. O. Finnemann, P. V. Christiansen, and P. Voetmann, Eds. Århus University Press, 2000, pp. 322–348.
- [13] A. Kubík, “Toward a formalization of emergence,” *Artificial Life*, vol. 9, pp. 41–65, 2003.
- [14] M. Resnick, *Turtles, Termites, and Traffic Jams*. MIT Press, 1994.
- [15] G. Theraulaz, E. Bonabeau, S. C. Nicolis, R. V. Solé, V. Fourcassié, S. Blanco, R. Fournier, J.-L. Joly, P. Fernández, A. Grimal, P. Dalle, and J.-L. Deneubourg, “Spatial patterns in ant colonies,” *Proc. Natl. Acad. Sci. USA*, vol. 99, no. 15, pp. 9645–9649, Jul. 2002.
- [16] A. Gajardo, A. Moreira, and E. Goles, “Complexity of Langton’s ant,” *Discrete Applied Mathematics*, vol. 117, pp. 41–50, 2002.
- [17] P. Dorbec and A. Gajardo, “Langton’s flies,” *Journal of Physics A: Mathematical and Theoretical*, vol. 41, no. 40, p. 405101, 2008.
- [18] H. Hamann, “Definition and behavior of Langton’s ant in three dimensions,” *Complex Systems*, vol. 14, no. 3, pp. 263–268, 2003.
- [19] L. Boltzmann, “Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie,” *Wien. Ber.*, vol. 53, 1866.
- [20] G. A. Gottwald and M. Oliver, “Boltzmann’s dilemma - an introduction to statistical mechanics via the Kac ring,” *SIAM Review*, vol. 51, pp. 613–635, 2009.
- [21] M. J. A. M. Brummelhuis and H. J. Hilhorts, “How a random walk covers a finite lattice,” *Physica A: Statistical Mechanics and its Applications*, vol. 185, no. 1–4, pp. 35–44, June 1992.
- [22] A. M. Nemirovsky, H. O. Martín, and M. D. Coutinho-Filho, “Universality in the lattice-covering time problem,” *Phys. Rev. A*, vol. 41, no. 2, p. 761767, August 1989.
- [23] H. Haken, *Synergetics - an introduction*. Berlin, Germany: Springer-Verlag, 1977.
- [24] H. Hamann, *Space-Time Continuous Models of Swarm Robotics Systems: Supporting Global-to-Local Programming*. Springer-Verlag, 2010.
- [25] J. C. Alexander, B. Giesen, R. Münch, and N. J. Smelser, Eds., *The Micro-Macro Link*. University of California Press, 1987.
- [26] M. Schillo, K. Fischer, and C. T. Klein, “The micro-macro link in DAI and sociology,” in *Multi-Agent-Based Simulation: Second International Workshop, Boston, MA, USA (MABS 2000)*, ser. LNCS, S. Moss and P. Davidsson, Eds., vol. 1979. Berlin, Germany: Springer-Verlag, 2000, pp. 303–317.
- [27] D. Yamins, “A theory of local-to-global algorithms for one-dimensional spatial multi-agent systems,” Ph.D. dissertation, Harvard University, Nov. 2007.
- [28] —, “Towards a theory of ‘local to global’ in distributed multi-agent systems,” in *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems (AAMS’05), Utrecht, Netherlands*, 2005, pp. 183–190.
- [29] R. Abbott, “Emergence explained: getting epiphenomena to do real work,” arXiv:cs/0602045, Feb. 2006.
- [30] D. A. Wolf-Gladrow, *Lattice-gas cellular automata and lattice Boltzmann models: an introduction*. Springer-Verlag, 2000.
- [31] F. J. Alexander, H. Chen, S. Chen, and G. D. Doolen, “Lattice boltzmann model for compressible fluids,” *Phys. Rev. A*, vol. 46, no. 4, pp. 1967–1970, Aug. 1992.
- [32] M. Kac, *Probability and related topics in physical sciences*. AMS Bookstore, 1957.
- [33] J. Loschmidt, “Über den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft, 1. Teil,” *Sitzungsber. Kais. Akad. Wiss. Wien, Math. Naturwiss.*, vol. 73, pp. 128–142, 1876.
- [34] E. W. Weisstein, “Langton’s Ant. From MathWorld—A Wolfram web resource,” <http://mathworld.wolfram.com/LangtonsAnt.html>, 2009.
- [35] L. Boltzmann, “Zu Hr. Zermelo’s Abhandlung ‘Ueber die mechanische Erklärung irreversibler Vorgänge’,” *Annalen der Physik*, vol. 60, pp. 392–398, 1897.
- [36] T. Schmickl, R. Thenius, C. Möslinger, G. Radspieler, S. Kernbach, and K. Crailsheim, “Get in touch: Cooperative decision making based on robot-to-robot collisions,” *Autonomous Agents and Multi-Agent Systems*, vol. 18, no. 1, pp. 133–155, February 2008.
- [37] R. Landauer, “Irreversibility and heat generation in the computing process,” *IBM Journal of Research and Development*, vol. 5, pp. 183–191, July 1961.
- [38] D. Polani, “Foundations and formalizations of self-organization,” in *Advances in Applied Self-organizing Systems*, ser. Advanced Information and Knowledge Processing, M. Prokopenko, Ed. Springer-Verlag, 2008.
- [39] R. P. Feynman, T. Hey, R. W. Allen, and T. Hey, *Feynman Lectures on Computation*. Cambridge, MA, USA: Perseus Books Group, 2000.
- [40] E. Fredkin and T. Toffoli, “Conservative logic,” *International Journal of Theoretical Physics*, vol. 21, p. 219253, 1982.
- [41] D. F. Wallace, *Everything and More: A Compact History of Infinity*. W. W. Norton & Company, 2003.
- [42] E. Mach, *Die Mechanik in ihrer Entstehung*, 3rd ed. Leipzig, Germany: Brockhaus, 1897.
- [43] M. A. Bedau, “Is weak emergence just in the mind?” *Minds & Machines*, vol. 18, pp. 443–459, 2008.
- [44] E. Zermelo, “Ueber einen Satz der Dynamik und die mechanische Wärmetheorie,” *Annalen der Physik*, vol. 57, pp. 485–494, 1896.
- [45] H. Poincaré, “Sur le problème des trois corps et les équations de la dynamique,” *Acta. Math.*, vol. 13, 1890.
- [46] L. Boltzmann, “Entgegnung auf die wärmetheoretischen Betrachtungen des Hr. E. Zermelo,” *Annalen der Physik*, vol. 57, pp. 773–784, 1896.
- [47] E. L. Post, “A variant of a recursively unsolvable problem,” *J. Symbolic Logic*, vol. 12, no. 2, pp. 255–56, 1946.
- [48] R. Groß, S. Magnenat, L. Küchler, V. Massaras, M. Bonani, and F. Mondada, “Towards an autonomous evolution of non-biological physical organisms,” in *Proc. of the 10th European Conf. on Artificial Life (ECAL’09)*, 2011, in press.
- [49] M. Gu, C. Weedbrook, Á. Perales, and M. A. Nielsen, “More really is different,” *Physica D*, vol. 238, pp. 835–839, 2009.